

Registration and entire shape acquisition for grid based active one-shot scanning techniques

Hiroshi Kawasaki, Takuto Hirukawa
Department of Information and Biomedical Engineering,
Kagoshima University, Kagoshima, Japan

Ryo Furukawa
The Graduate School of Information Science,
Hiroshima City University, Hiroshima, Japan

Abstract—One-shot active stereo using structured light is a practical solution for dynamic scene acquisition. Basically, those methods are based on encoding positional information of the pixel into the single projected pattern. A disadvantage of such methods is decreases of the spatial resolution caused by requiring a certain area of the pattern to encode the positional information. Among those methods, grid-based patterns are promising at the point of accuracy and robustness, since triangulation for 3D reconstruction is conducted with light-sectioning method and a line detection is usually a stable image processing. However, no shapes are recovered between the grid lines, and thus, the whole reconstructed shape tends to be sparse. To deal with the problem, integrating multiple shapes that are sequentially captured using registration algorithm such as ICP is one solution. In previous work, we show that naive ICP works poorly for grid-like structured point clouds, and proposed a specialized ICP algorithm for aligning a set of grid-like structured 3D shapes. In this paper, we extend this approach and propose a process for entire shape modeling by capturing objects from all the directions using turn table, and integrating into a single shape using our improved ICP. To achieve this, setting good initial 3D shapes is important. For solution, we interpolation grid shapes to create smooth surface so that common ICP works. Comprehensive experiments are conducted to show the strength of our method compared to common ICP.

I. INTRODUCTION

For visual information of robots, analysis and inspection purposes of human motion, there is a high demand to capture 3D shapes of dynamic scenes. To capture dynamic scenes, one-shot active stereo methods using structured illuminations are practical solutions, and widely researched and developed [1]. Basically, those methods are based on encoding positional information of each point on the projected image into the pattern of the image. To achieve 3D reconstruction, the signals of the captured camera image should be correctly decoded to obtain the positional information on the projector image for acquiring correspondences between the projector image plane and the camera image plane. A disadvantage of such methods is decreases of the spatial resolution caused by assuming a certain area of smooth surfaces to encode the positional information with sufficient uniqueness of the encoded information.

Among the one-shot active stereo methods, grid based reconstruction methods have an advantage of accurate positioning of reconstructed points [2], [3], [4], because triangulation for obtaining 3D points is done by light-sectioning method, where 3D reconstruction of sub-pixel accuracy can be easily

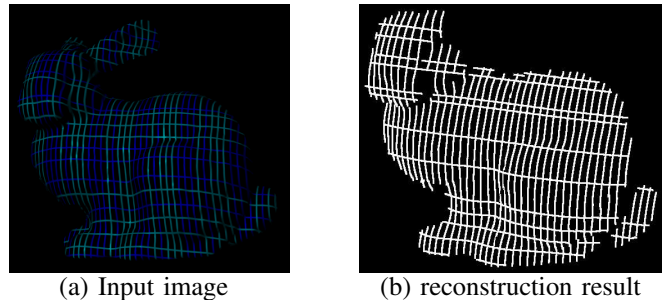


Fig. 1. Grid reconstruction result of Stanford bunny

achieved. Moreover, detection of line patterns is generally stable even with noises such as subsurface scattering effects.

One disadvantage of 3D capturing method using the grid patterns is sparse results obtained, because the intervals of the grid lines remains empty (Fig. 1). To deal with the problem, integrating multiple shapes that are sequentially captured using registration algorithm such as ICP is promising. In [5], we proposed an alignment method for 3D data set acquired by a 3D endoscopic system, where grid-like structured light is used. In the work, multiple shapes that consist of 3D curves are aligned with specially-designed alignment criteria, obtaining dense 3D points from sparse, line-based shapes. Note that they assume mainly small translational motion because their main purpose is to increase the density of the shape rather than enlarge the recovered shape. In this work, we extend the method to realize entire-shape acquisition using an active scanning method using grid-based structured light.

In this paper, we show that naive ICP works poorly for grid-like structured point clouds, and show an improved ICP algorithm [5] that takes grid-like structures of the inputs into consideration. Using the improved algorithm, multiple grid-structured point clouds can be registered correctly. We also newly propose a process for entire shape modeling by capturing objects from all the directions using grid-pattern projection. This is efficiently done by densifying the point clouds to realize smoothness on the surface, whereas original shapes only have high frequency structure, *i.e.*, grid lines. In the proposed modeling process, our improved ICP is utilized for registering multiple grid-like shapes into an integrated shape model. In the experiments, we show the effectiveness of our technique with several tests using simulated and real

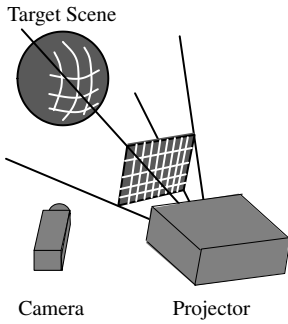


Fig. 2. System configuration

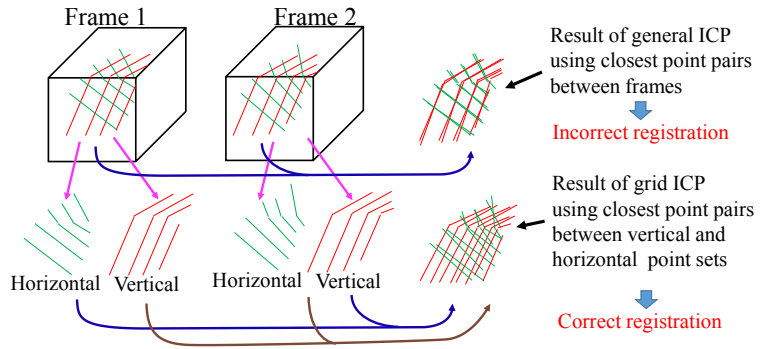


Fig. 3. Process of grid ICP

data by comparing with common ICP method.

II. RELATED WORK

Our contribution is based on a rigid registration algorithm. Rigid registration algorithms estimate translation and rotation of an object from two point sets, with the ICP algorithm [6] and its extension to multiple point sets [7] being the two best-known approaches. Since then, improved techniques have been intensively researched on realtime registration [8], large scale simultaneous registration [9], [10] and color compensated registration [11]. However, since they all assume a large overlap of dense shapes, they generally cannot be used whenever the shape is sparse such as a grid based reconstruction [2], [3], [12].

Recently, an ICP for the sparse point set was proposed [13], however, since the technique is still based on the correspondences of closest points, lines in the same direction are inevitably pulled together, and thus, all the grid based shapes are bundled into a single grid liked shape.

Banno *et al.* proposed a method to align the multiple 3D curves which are reconstructed by the light sectioning method into single consistent shape [14], however, they assumed that a base dense shape with holes is captured in advance as the target object for aligning 3D curves. Therefore, the technique cannot be applied to the data which consists of independent curves only. Another approach to achieving robust registration of multiple shapes is based on 3D features extracted from input shapes [15], [16], [17], [18], [19]. However, stable 3D features are usually extracted only from dense 3D points and cannot be applied to grid based shapes, whose points are sparse and unevenly distributed.

The proposed method is an extension of the alignment method of [5]. In [5], an alignment method is used for making dense point clouds from a set of sparse point sets that consist of 3D lines. In this work, we use the similar approach for realizing an entire-shape modeling. To achieve this, setting good initial estimation of the relative 3D positions of the multiple 3D shapes is important. In this paper, we propose a novel method for this problem.

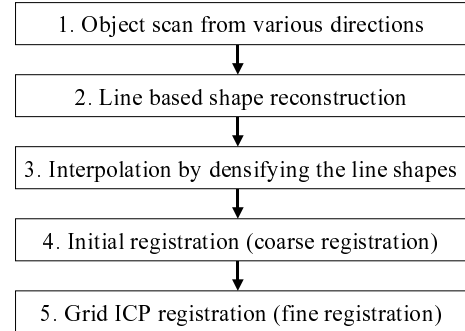


Fig. 4. Overview of the registration of grid-line shapes.

III. REGISTRATION OF GRID-STRUCTURED SHAPES AND ENTIRE SHAPE ACQUISITION BY GRID-PATTERN PROJECTION

Overview of our technique is shown in Fig. 4. First, a target object is captured from various directions by line based scanning system, especially a grid based stereo system using a video projector [2], [3], [12] in the paper. We used a rotation turn table in our experiment for scanning (Sec. III-A). To recover the entire shape, registration of all the shapes is required, however, since they are thin (line) shapes, common registration techniques do not work; note that it is difficult to either estimate normal direction of points on a 3D line/curve or find correct correspondences between lines/curves. To avoid the problem, we first densify the shapes so that a common registration algorithm, such as ICP, can be applied (Sec. III-B). After the initial registration, we apply grid ICP which is specialized for grid shapes to achieve fine registration (Sec. III-C).

A. Active stereo system with grid-pattern projection

A sequence of 3D point clouds is captured by a single pattern projector and a camera as shown in Fig. 2. The camera and the projector are assumed to be calibrated (*i.e.*, the intrinsic and extrinsic parameters are known). Since the projector casts a static grid pattern, no synchronization is required and it is suitable to acquire a 3D shape of dynamic scene or entire shape by moving the sensor. To recover the shape from single projection, correspondences between the detected line on the captured image and the line on the projection pattern

are estimated [2], [3], [12]. Once the correspondences are obtained, the points on the vertical and horizontal lines are reconstructed as 3D curves using a light-sectioning method. Since the line intervals between parallel lines are wide enough to avoid mis-detection by the subsurface scattering or similar effect, the shapes can be only coarsely reconstructed. To increase the density of the sparse grid shaped 3D points, one solution is to capture the object multiple times by moving the sensor, and then, align to integrate them.

B. Reconstruction of entire shapes of objects

To achieve acquiring an entire shape of object, a large number of images are taken from various directions. For practicality of capturing process, usually objects are captured by 3D sensor, rotating it with a turntable at most 20 to 40 times. Initial position can be roughly estimated by using the marker on the turntable or image/shape features on the object. In the experiments, we calibrate the turntable to estimate the rotation axis with respect to the camera in advance to the capturing process and

One critical issue of the scenario is that, since ICP algorithm assumes smooth surface of the object, simultaneous registration of high frequency shape with a large number of input easily stack into local minimum; remember that a sparse grid shape is a considerably high frequency shape.

To deal with this problem, we interpolate the sparse grid shapes to make dense and smooth surface. The interpolation is done in the 2D domain of the depth image of the camera, *i.e.*, the depth function $d(\mathbf{x}) : \mathcal{R}^2 \rightarrow \mathcal{R}$ is interpolated from the sample depth values $s(\mathbf{x}_i) \in \mathcal{R}$ at $\mathbf{x}_i \in \mathcal{R}^2$. We use approximation form of

$$d(\mathbf{x}) = \frac{\sum_i \phi(\|\mathbf{x} - \mathbf{x}_i\|)g_i(\mathbf{x})}{\sum_i \phi(\|\mathbf{x} - \mathbf{x}_i\|)}, \quad (1)$$

where $g_i(\mathbf{x})$ is a linear approximation of $d(\mathbf{x})$ around the sample point \mathbf{x}_i . $\phi(\|\mathbf{x} - \mathbf{x}_i\|)$ is a value for weighting $p_i(\mathbf{x})$ most at \mathbf{x}_i and less at \mathbf{x} that is far from \mathbf{x}_i . For the weighting function ϕ , we use $\phi(t) = \exp(-\frac{t^2}{R_w^2})$ where R_w is the radius of approximation.

For the approximation $g_i(\mathbf{x})$, we use linear 2D regression of depth image $d(\mathbf{x})$ around \mathbf{x}_i , which is fit to the sample set around \mathbf{x}_i , *i.e.*, $N(\mathbf{x}_i) = \{(\mathbf{x}_j, s(\mathbf{x}_j)) \mid \|\mathbf{x}_i - \mathbf{x}_j\| < R_p\}$, where R_p is the radius of the neighbor samples. The regression model is $g_i(\mathbf{x}) = \mathbf{a}_i \cdot (\mathbf{x} - \mathbf{x}_i) + s(\mathbf{x}_i)$ where \mathbf{a}_i is the liner regression coefficient calculated from $N(\mathbf{x}_i)$, *i.e.*, $\mathbf{a}_i = (\sum_{\mathbf{x}_j \in N(\mathbf{x}_i)} (\mathbf{x}_j - \mathbf{x}_i)(\mathbf{x}_j - \mathbf{x}_i)^T)^{-1} (\sum_{\mathbf{x}_j \in N(\mathbf{x}_i)} (\mathbf{x}_j - \mathbf{x}_i)(s(\mathbf{x}_j) - s(\mathbf{x}_i)))$.

The interpolated shape may miss the small details of the original shape, but they are useful for coarse and robust registration of the shapes. Therefore, we first register these interpolated shapes using ICP, and then, apply the estimated rigid transformations to each original grid shapes, respectively.

C. Grid ICP algorithm for grid-structured shapes

Once the coarse registration is achieved with the interpolated shapes, fine registration is applied to the original shapes. Although the ICP algorithm is the most used solution to

conduct shape alignment between 3D shapes of a static object, we cannot simply use the algorithm to the line shaped object. The common ICP algorithm consists of two steps such as 1) searching for the closest point $\mathbf{q}_i \in \mathcal{R}^3$ of the scene object from point $\mathbf{p}_i \in \mathcal{R}^3$, which belongs to the target object, and 2) estimate a rigid transformation R, \mathbf{t} by minimizing

$$\sum_i \|\mathbf{p}_i - (R \mathbf{q}_i + \mathbf{t})\|^2. \quad (2)$$

Final parameters of R, \mathbf{t} are obtained by iterating the two steps until convergence.

The reason why such a naive ICP algorithm does not work properly on sparse grid shapes is that the closest points from vertical/horizontal lines of the scene object are usually found on the line in same direction of the target shape; note that such incorrect corresponding points are pulled together to minimize the differences to configure an incorrect wrong shape as shown in Fig. 3 upper row. Noteworthy, if multiple shapes are captured with small shifts (translational motions) of positions, grid lines tends to be bundled together.

For solution, we use a new ICP algorithm, which we call grid ICP algorithm in this paper, to solve this problem. Fig. 3 lower row shows the process of our algorithm. We first divide the grid shape into two sets of lines depending on the line directions, *i.e.*, the vertical set and the horizontal set. Then, the closest point \mathbf{q}_i^v in the vertical line set from the point \mathbf{p}_i^h in the horizontal line set is searched. We use KD-tree to find correspondences in the paper. Similarly, the closest point \mathbf{q}_i^h in the horizontal line set from the point \mathbf{p}_i^v which belongs to the vertical line set is found. Finally, rigid transformation parameters R, \mathbf{t} are estimated by minimizing

$$\sum_i \|\mathbf{p}_i^h - (R \mathbf{q}_i^v + \mathbf{t})\|^2 + \sum_j \|\mathbf{p}_j^v - (R \mathbf{q}_j^h + \mathbf{t})\|^2. \quad (3)$$

Final results are obtained by iterating the aforementioned steps until convergence. Within this scenario, grid lines of the final shapes are evenly distributed realizing dense reconstruction of the object surface.

IV. EXPERIMENTS

A. Entire-shape acquisition using simulation data

To evaluate the entire-shape acquisition, first, we made a simulated data. The model for the simulation is Stanford bunny shape data. The projected pattern is a straight-lined grid pattern with modulated intervals used in [3]. The image with the pattern projection by a virtual projector was generated by a simple computer graphics technique. To simulate the entire-shape acquisition, a number of images were similarly generated while rotating the rendered shape model. By processing each of the rendered image with the 3D reconstruction process for the pattern [3], grid-structured shapes of 3D curves were obtained as shown in Fig. 5(a)-(d). Then, the entire-shape modeling process of the smooth shape interpolation method followed by the grid ICP is applied to 8 of them. Fig. 5(i) shows densified shapes at initial positions, and Fig. 5(j) shows a result after ICP as for the coarse registration. A common ICP

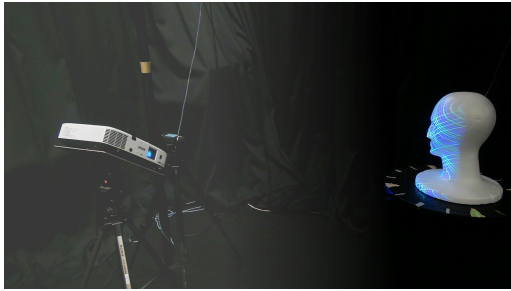


Fig. 6. Entire object capturing scene.

and grid ICP are applied on Fig. 5(j). As shown in Fig. 5(l), grid ICP method correctly recovered the entire shape, whereas result of common ICP, as shown in Fig. 5(k), was stuck into the local minimum and the shape is shrank and distorted. RMS errors are 2.73% after ICP on interpolated shape (coarse registration), 2.71% after grid ICP and 6.03% with common ICP of the height of the target object. Note that common ICP result is even worse than the coarse ICP result for this case.

B. Evaluation of entire-shape acquisition using real objects

Next, we evaluated grid ICP technique for translational movements of the target object, *i.e.*, mannequin head, by using a video projector and a CCD camera for the experiment and the actual captured scene is shown in Fig. 6. We use grid reconstruction method of Sagawa *et al.* [3] by which accurate 3D shapes of the texture-less objects can be stably obtained by using color information. We captured 9 sequential frames of images of the target objects while slightly moving the capturing device of projector and camera pair with $3 * 3$ shifts to x and y direction, respectively. The captured images are reconstructed frame-by-frame, and a common ICP and the proposed grid ICP are applied to the shapes for comparison. Fig. 7(c)(d) are examples of a reconstructed shapes from the captured frames. Fig. 7 (a) is the 3D shape reconstructed with a temporal encoded technique, *i.e.*, gray code [20] as the ground truth. The registration result with a common ICP and the grid ICP are shown in Fig. 7(e)(f). From the figures, we can confirm that integrated 3D points with the grid ICP are more evenly distributed than that with the common ICP.

Then, we captured the entire shape of the mannequin head using the rotation table with the same system. We captured 32 times and apply our technique to 8 out of them. Fig. 7(g)(h) shows examples of a reconstructed shapes from the captured frames. Those reconstructed frames are registered with both the common ICP and the grid ICP algorithms and the registration results are shown in Fig. 7(i)(j). From the figures, we can confirm that the entire shape is correctly reconstructed with the techniques. As can be seen in Fig. 7(i)(j), grid lines are more sparse with the common ICP than that with the grid ICP algorithm, since the grid lines of multiple frames were pulled together with a common ICP, whereas the grid lines were uniformly distributed with the grid ICP.

Finally, we conducted the same experiments to several objects and all results are shown in Fig. 9. The integrated

shapes are registered to the ground-truth shape (Fig. 9 top row) and RMSE values are calculated for both the common ICP and the grid ICP algorithms. The RMSE values are summarized in Table. I, II, Fig. 8(a) and (b). From the data, we can confirm that the grid ICP always outperforms the others both in translational and rotational movements. At the same time, we can also confirm that the ratio of the improvement is greater in translational movement than in rotational movement, *i.e.*, 6.6% for translation and 3.0% for rotation in average. Such differences can be intuitively understood by the fact that the translational movement does not drastically change the recovered shape, whereas rotational movement makes a large difference, if the shape is not planar.

V. CONCLUSION

In this paper, we proposed an improved ICP algorithm specifically for registering multiple grid-structured point clouds to recover a entire shape of object. The idea of the proposed algorithm is based on observation that the naive ICP often fails for grid-structured shapes because vertical lines and vertical lines (or horizontal lines and horizontal lines) tend to improperly attract each other. In the proposed algorithm, the cost function to be minimized is defined so that the vertical lines only attract horizontal lines, and vice versa. We also proposed a coarse to fine technique, in which thin shapes are densified to create smooth surface so that common ICP works with rough initial positions. The potential of the techniques was verified by experiments using simulated and real data. Our future work is to restoring small detailed shapes from relatively coarse grid-pattern projection.

REFERENCES

- [1] Microsoft, "Xbox 360 Kinect," 2010, <http://www.xbox.com/en-US/kinect>.
- [2] H. Kawasaki, R. Furukawa, R. Sagawa, and Y. Yagi, "Dynamic scene shape reconstruction using a single structured light pattern," in *CVPR*, June 23-28 2008, pp. 1-8.
- [3] R. Sagawa, Y. Ota, Y. Yagi, R. Furukawa, N. Asada, and H. Kawasaki, "Dense 3D reconstruction method using a single pattern for fast moving object," in *ICCV*, 2009.
- [4] A. O. Ulusoy, F. Calakli, and G. Taubin, "One-shot scanning using debruijn spaced grids," in *The 7th IEEE Conf. 3DIM*, 2009, pp. 1786-1792.
- [5] R. Furukawa, H. Morinaga, Y. Sanomura, S. Tanaka, S. Yoshida, and H. Kawasaki, "Shape acquisition and registration for 3d endoscope based on grid pattern projection," in *ECCV2016(to appear)*, 2016.
- [6] P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239-256, 1992.
- [7] P. NEUGEBAUER, "Geometrical cloning of 3d objects via simultaneous registration of multiple range image," *Proc. Int'l Conf. Shape Modeling and Applications*, pp. 130-139, 1997. [Online]. Available: <http://ci.nii.ac.jp/naid/10008212880/>
- [8] O. Hall-Holt and S. Rusinkiewicz, "Stripe boundary codes for real-time structured-light range scanning of moving objects," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 2. IEEE, 2001, pp. 359-366.
- [9] T. Ooishi, R. Sagawa, A. Nakazawa, R. Kurazume, and K. Ikeuchi, "Parallel alignment of a large number of range images," in *IEEE Conf. 3DIM2003*, 2003, pp. 195-202.
- [10] T. Oishi, R. Kurazume, A. Nakazawa, and K. Ikeuchi, "Fast simultaneous alignment of multiple range images using index images," in *3-D Digital Imaging and Modeling, 2005. 3DIM 2005. Fifth International Conference on*. IEEE, 2005, pp. 476-483.

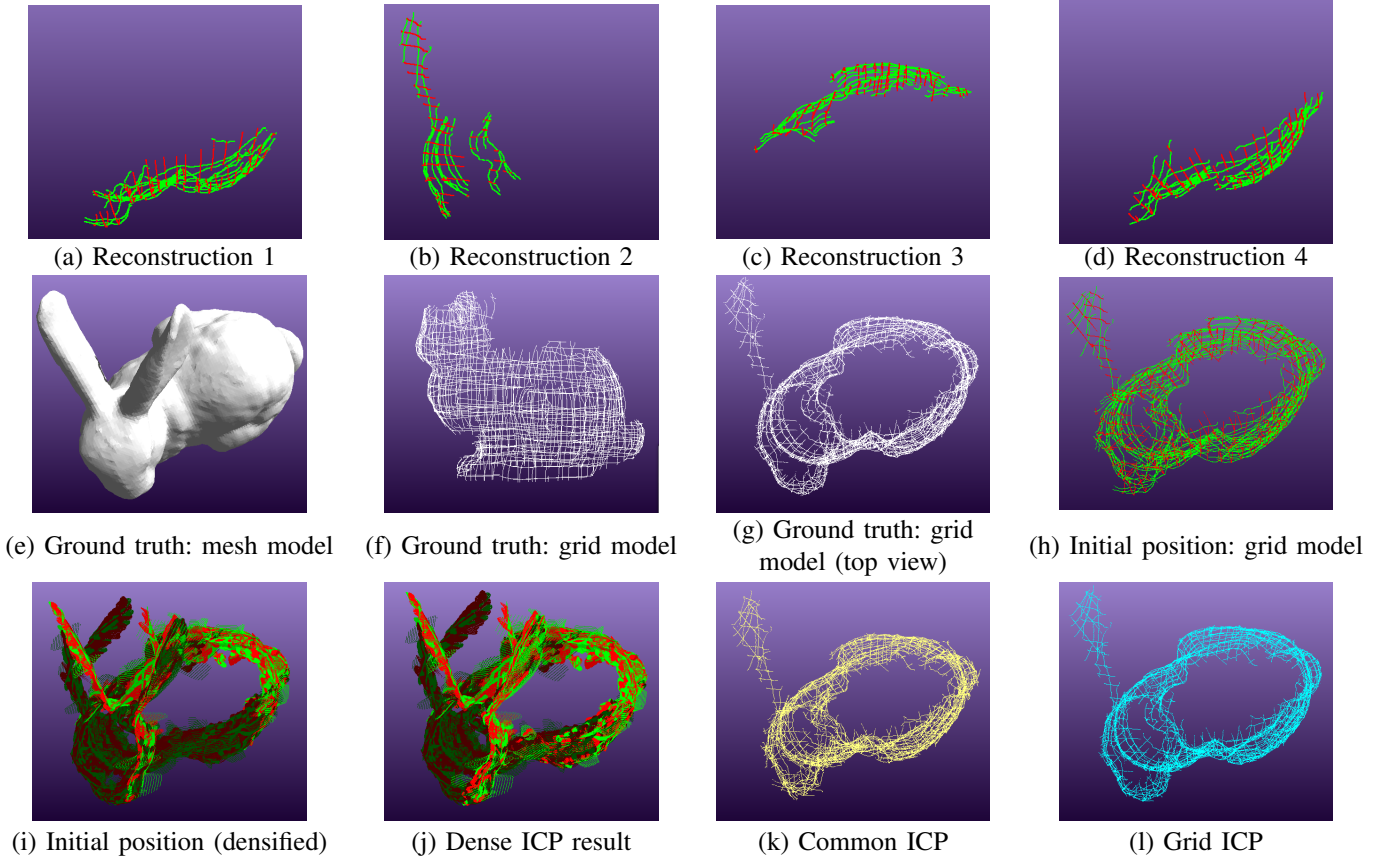


Fig. 5. Bunny object is reconstructed with the grid scan technique [2]. Random noise is added on the shapes and initial positions. Coarse registration is first applied, and then, common ICP and grid ICP algorithms are applied. We can clearly see that grid ICP recovers its correct positions, whereas common ICP shrinks the shapes.

TABLE I
RMSE OF REGISTRATION WITH TRANSLATIONAL MOVEMENT.

	Owl	Boots	Pyramid	Mannequin	Mars	Skull
Common ICP (mm)	1.115	2.494	0.679	0.939	1.013	1.167
Grid ICP (mm)	0.929	2.384	0.628	0.892	0.966	1.147

TABLE II
RMSE OF REGISTRATION WITH ROTATIONAL MOVEMENT.

	Owl	Boots	Pyramid	Mannequin	Mars	Skull
Coarse registration (mm)	1.497	2.397	0.780	1.545	1.840	1.685
Common ICP (mm)	1.221	2.302	0.821	1.352	1.758	1.443
Grid ICP (mm)	1.169	2.271	0.742	1.350	1.758	1.404

- [11] A. E. Johnson and S. B. Kang, "Registration and integration of textured 3d data," *Image and vision computing*, vol. 17, no. 2, pp. 135–147, 1999.
- [12] R. Sagawa, H. Kawasaki, R. Furukawa, and S. Kiyota, "Dense one-shot 3D reconstruction by detecting continuous regions with parallel line projection," in *ICCV*, 2011, pp. 1911–1918.
- [13] S. Bouaziz, A. Tagliasacchi, and M. Pauly, "Sparse iterative closest point," *Computer Graphics Forum (Symposium on Geometry Processing)*, vol. 32, no. 5, pp. 1–11, 2013.
- [14] A. Banno, T. Masuda, T. Oishi, and K. Ikeuchi, "Flying laser range sensor for large-scale site-modeling and its applications in bayon digital archival project," *Int. J. Comput. Vision*, vol. 78, no. 2-3, pp. 207–222, Jul. 2008. [Online]. Available: <http://dx.doi.org/10.1007/s11263-007-0104-6>
- [15] H. Li and R. Hartley, "The 3d-3d registration problem revisited," in *Proc. Int'l Conf. Computer Vision*, 2007, pp. 1–8.
- [16] J. Yang, H. Li, and Y. Jia, "Go-icp: Solving 3d registration efficiently and globally optimally," in *IEEE Int'l Conf. Computer Vision*, 2013, pp. 1457–1464.
- [17] R. Wang, J. Choi, and G. Medioni, "3d modeling from wide baseline range scans using contour coherence," *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4018–4025, June 2014.
- [18] R. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (FPFH) for 3d registration," in *Int'l Conf. Robotics and Automation*, May 2009, pp. 3212–3217.
- [19] R. Rusu, N. Blodow, Z. Marton, and M. Beetz, "Aligning point cloud views using persistent feature histograms," in *Int'l Conf. Intelligent Robots and Systems*, 2008, pp. 3384–3391.
- [20] K. Sato and S. Inokuchi, "Range-imaging system utilizing nematic liquid crystal mask," in *Proc. Int. Conf. on Computer Vision*, 1987, pp. 657–661.
- [21] R. Furukawa, M. Aoyama, S. Hiura, H. Aoki, Y. Kominami, Y. Sanomura, S. Yoshida, S. Tanaka, R. Sagawa, and H. Kawasaki, "Calibration of a 3d endoscopic system based on active stereo method for shape measurement of biological tissues and specimen," in *EMBC*, 2014, pp. 4991–4994.
- [22] R. Furukawa, R. Masutani, D. Miyazaki, M. Baba, S. Hiura, M. Visentini-Scarzanella, H. Morinaga, H. Kawasaki, and R. Sagawa, "2-dof auto-calibration for a 3d endoscope system based on active stereo," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, Aug 2015, pp. 7937–7941.

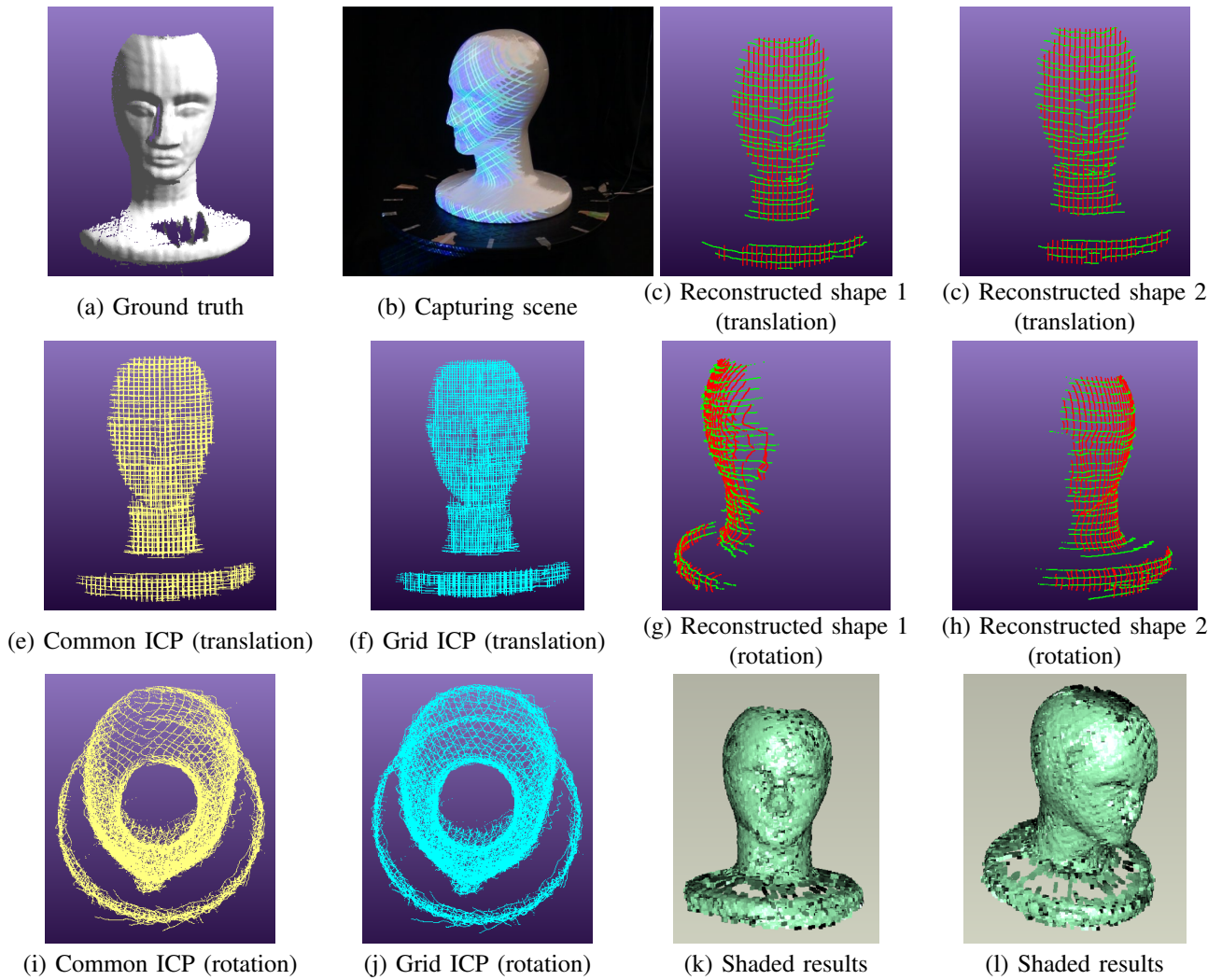


Fig. 7. Mannequin heads are reconstructed by the grid scan technique [2] from several camera positions. The common ICP is applied to the densified shapes as a coarse registration, and then, common ICP and grid ICP algorithms are applied to the grid-line shapes. Results are shown that shapes are more evenly distributed with the grid ICP than the common ICP techniques for both translational and rotational motions.

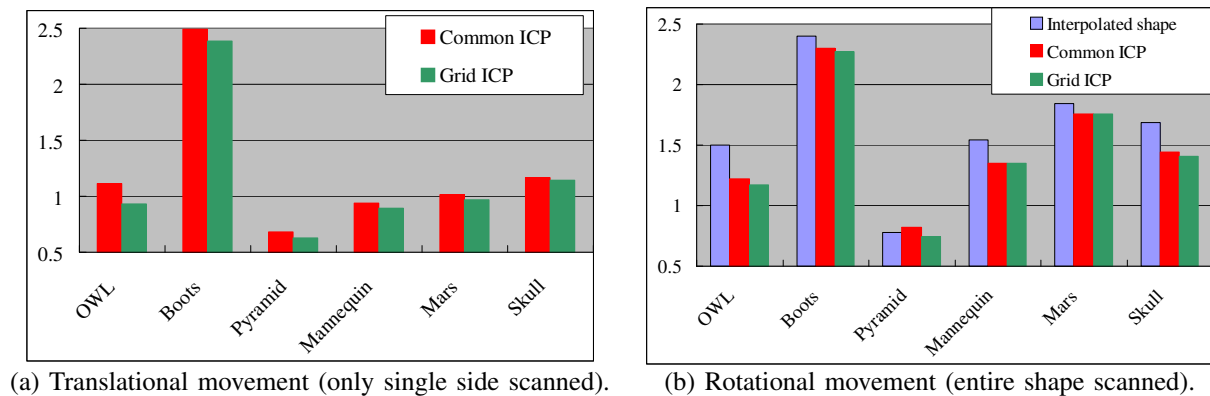


Fig. 8. Registration comparison results. RMSE (mm) values are minimum for all the cases with grid ICP technique.

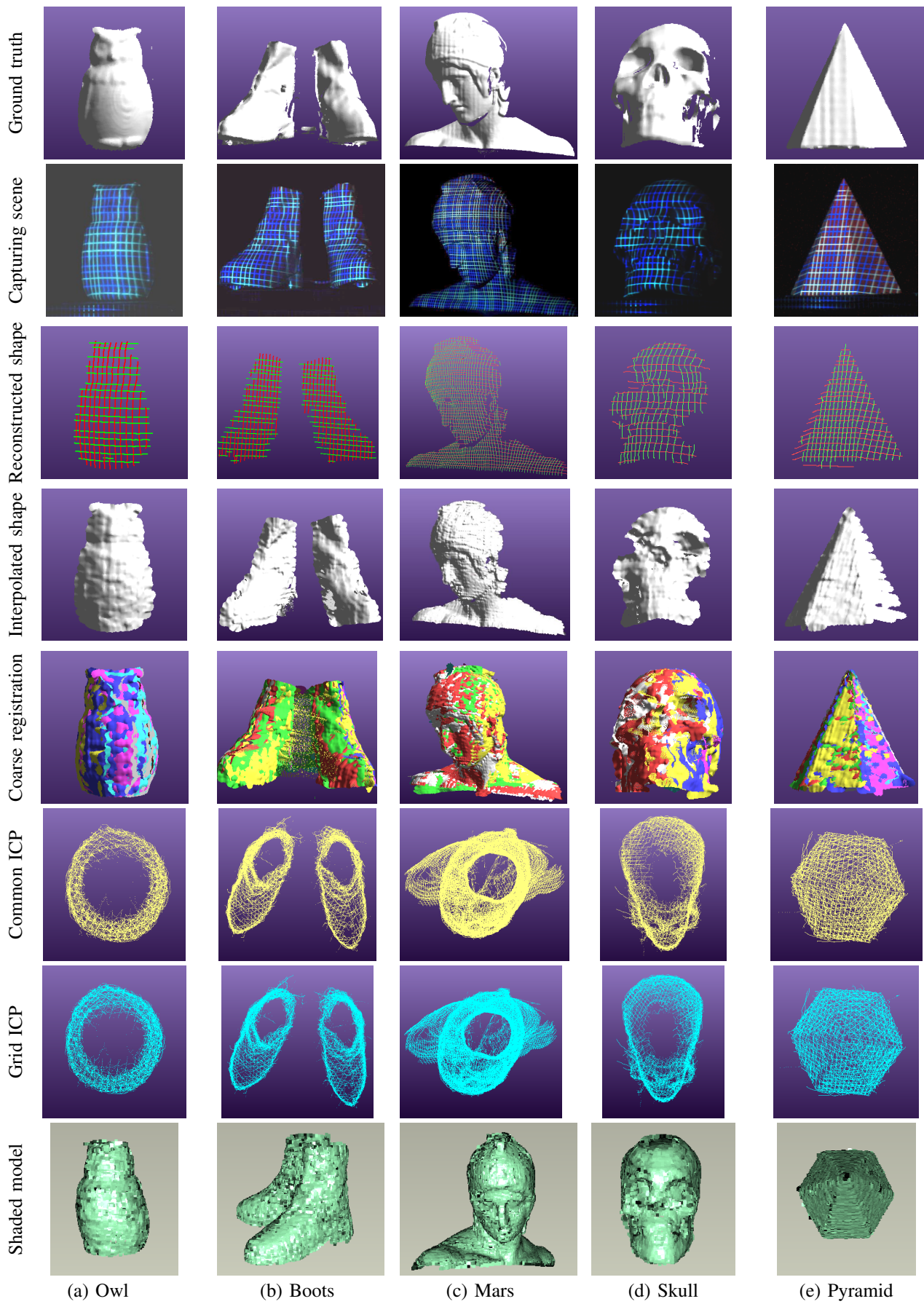


Fig. 9. Entire shape registration results.